BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

1/3/1
DIALOG(R)File 351:Derwent WPI

010516273 **Image available**
WPI Acc No: 1996-013224/ **199602**
XRPX Acc No: N96-011349

Communication control method of parallel processor system - involves performing course selection of network based on physical address

Patent Assignee: HITACHI LTD (HITA)
Inventor: SUZUKI K
Number of Countries: 002 Number of Patents: 003
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 7262146	A	19951013	JP 9447245	A	19940317	199602	B
US 5808886	A	19980915	US 95403998	A	19950314	199844	
JP 3402398	B2	20030506	JP 9447245	A	19940317	200330	

Priority Applications (No Type Date): JP 9447245 A 19940317
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 7262146	A		9	G06F-015/16	
US 5808886	A			G05B-019/18	
JP 3402398	B2		9	G06F-015/177	Previous Publ. patent JP 7262146

[redacted]

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-262146
(43)Date of publication of application : 13.10.1995

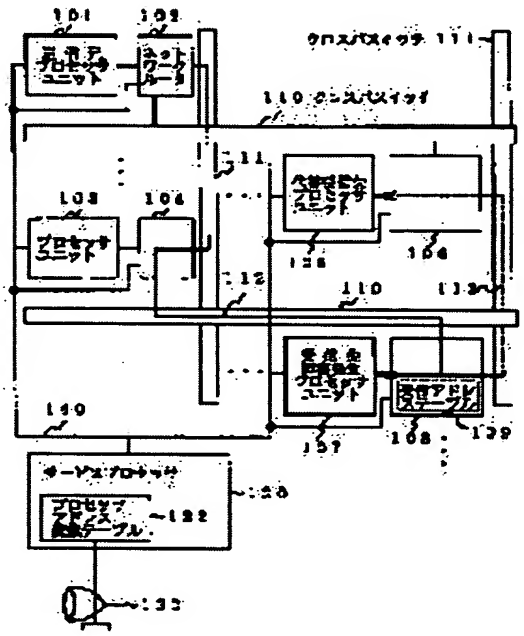
(51)Int.Cl. G06F 15/16
G06F 12/10
G06F 15/173

(21)Application number : 06-047245
(22)Date of filing : 17.03.1994

(71)Applicant : HITACHI LTD
(72)Inventor : SUZUKI KAORU

(54) COMMUNICATION CONTROL METHOD FOR PARALLEL PROCESSOR SYSTEM

(57)Abstract:
PURPOSE: To quickly and efficiently perform the dynamic reconstitution processing of packet network routing.
CONSTITUTION: A transmission source processor unit 101 adds the logical address of a reception destination process unit to a packet to transmit the packet. Network routers 102 to 108 refer to a processor address conversion table 122 through a signal line 140 to obtain a physical address corresponding to the destination logical address and set a route 112 to transfer the packet to the reception destination processor unit 107. If a fault occurs in the reception destination 107, a service processor 120 changes correspondence between logical and physical addresses in the conversion table 122. As the result, a route 113 to an alternative processor unit 105 is dynamically set.



LEGAL STATUS

[Date of request for examination]	15.07.1999
[Date of sending the examiner's decision of rejection]	
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	3402398
[Date of registration]	28.02.2003
[Number of appeal against examiner's decision of rejection]	
[Date of requesting appeal against examiner's decision of rejection]	
[Date of extinction of right]	

(19)日本国特許庁（ＪＰ）　(12)**公開特許公報（A）**　(11)特許出願公開番号

特開平7－262146

(43)公開日　平成７年(1995)10月13日

(51)Int.Cl.[6]		識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０６Ｆ	15/16	４７０	Ｈ		
	12/10		Ｋ　7608－5B		
	15/173				
				Ｇ０６Ｆ　15/ 16　４００　Ｖ	

審査請求　未請求　請求項の数４　ＯＬ　（全９頁）

(21)出願番号　特願平6－47245

(22)出願日　平成６年(1994)３月17日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　鈴木　薫
神奈川県秦野市堀山下１番地　株式会社日立製作所汎用コンピュータ株式会社内

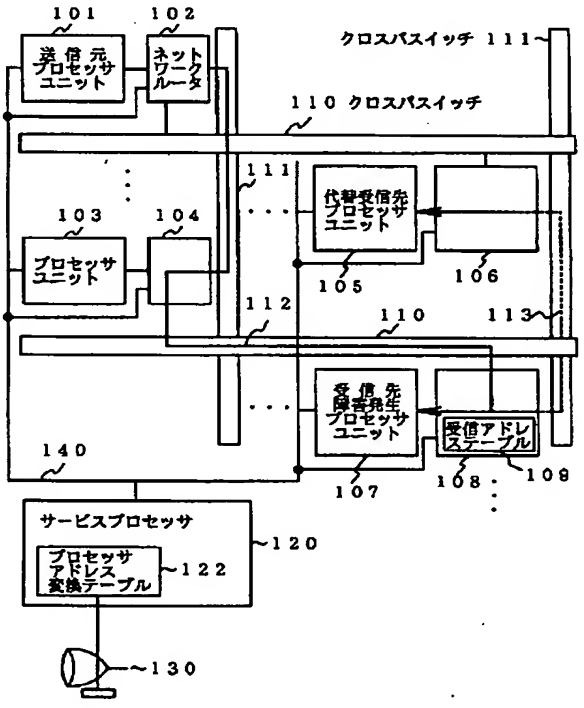
(74)代理人　弁理士　鈴木　誠

(54)【発明の名称】　並列プロセッサシステムの通信制御方法

(57)【要約】

【目的】　パケットのネットワークルーチングの動的再構成処理を高速かつ効率的に行う。

【構成】　送信元プロセッサユニット１０１は、受信先プロセッサユニットの論理アドレスを付加してパケットを送信する。ネットワークルータ１０２～１０８は、信号線１４０を介しプロセッサアドレス変換テーブル１２２を参照して宛先論理アドレスに対応する物理アドレスを得、ルート１１２を設定して、受信先プロセッサユニット１０７にパケットを転送する。受信先１０７で障害が発生すると、サービスプロセッサ１２０は変換テーブル１２２の論理／物理アドレスの対応を変更する。この結果、代替プロセッサユニット１０５へのルート１１３が動的に設定される。



1

【特許請求の範囲】

【請求項１】　複数のプロセッサユニットが各々ネットワークルータを介してネットワーク網で結合され、プロセッサユニット間のデータ転送がパケット送受信により行われる並列プロセッサシステムにおける通信制御方法であって、

前記複数のプロセッサユニット中の任意プロセッサユニットあるいは別に設けた、システムの状態を監視する監視プロセッサに、各プロセッサユニットの論理アドレスと物理アドレスとの対応を登録したプロセッサアドレス変換テーブルを設け、

パケット送信元プロセッサユニットは自プロセッサユニットアドレス及び受信先プロセッサユニットアドレスとして論理アドレスを付加したパケットを送信し、

前記パケットを受信したネットワークルータは、前記監視プロセッサに要求を出して受信先プロセッサユニットの論理アドレスに対応する物理アドレスを得、該物理アドレスにもとづいてネットワーク網の経路選択を行う、

ことを特徴とする並列プロセッサシステムの通信制御方法。

【請求項２】　複数のプロセッサユニットが各々ネットワークルータを介してネットワーク網で結合され、プロセッサユニット間のデータ転送がパケット送受信により行われる並列プロセッサシステムにおける通信制御方法であって、

各ネットワークルータに、該並列プロセッサシステムを構成する各プロセッサユニットの論理アドレスと物理アドレスとの対応を登録したプロセッサアドレス変換テーブルを設け、

パケット送信元プロセッサユニットは自プロセッサユニットアドレス及び受信先プロセッサユニットアドレスとして論理アドレスを付加したパケットを送信し、

前記パケットを受信したネットワークルータは、前記プロセッサアドレス変換テーブルを参照して受信先プロセッサユニットの論理アドレスに対応する物理アドレスを得、該物理アドレスにもとづいてネットワーク網の経路選択を行う、

ことを特徴とする並列プロセッサシステムの通信制御方法。

【請求項３】　請求項１もしくは２記載の並列プロセッサシステムの通信制御方法において、プロセッサアドレス変換テーブルの論理アドレスと物理アドレスの対応を変更して、ネットワーク網から任意プロセッサユニットを論理的に切り離すことを特徴とする並列プロセッサシステムの通信制御方法。

【請求項４】　請求項３記載の並列プロセッサシステムの通信制御方法において、

各ネットワークルータに当該ネットワークルータに接続されたプロセッサユニットがパケット受信処理中であるか否かを示す識別ビットと当該パケットの送信元プロセッサユニットの論理アドレスとの対応を登録する受信アドレステーブルを設け、

受信先プロセッサユニットの障害発生時、監視プロセッサが当該障害発生プロセッサユニットに接続されたネットワークルータの受信アドレステーブルを読み込み、前記識別１ビットがパケット受信処理中を示していれば、対応する論理アドレスのパケット送信元プロセッサユニットに対してパケット再送を通知することを特徴とする並列プロセッサシステムの通信制御方法。

2

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、複数のプロセッサユニットが各々ネットワークルータを介してネットワーク網で結合され、プロセッサユニット間のデータ転送がパケット送受信によって行われる並列プロセッサシステムにおける通信制御方法に係り、特にプロセッサユニットの動的再構成を効率良く行うのに好適な並列プロセッサシステムの通信制御方法に関する。

【０００２】

【従来の技術】単一プロセッサの高速化の限界と、大量のデータの高速処理要求に伴って、例えばＳＩＭＤ（Ｓingle Ｉnstruction Ｓtream Multiple Data Ｓtream；各プロセッサが異なるデータに対して、同一の命令を実行する方式）、ＭＩＭＤ（Multiple Ｉnstruction Ｓtream Multiple Data Ｓtream；各プロセッサが独立に個別の命令を実行する方式）など種々の方式の並列処理マシンが開発され、実用に供されている。通常、この種の並列プロセッサシステムは、数百から数万のプロセッサユニットで構成されている。

【０００３】従来、このような並列プロセッサシステムにおいて、プロセッサユニット間のデータ転送をパケット送受信によって行う場合、実際にプロセッサユニットが接続されている物理ネットワークアドレスを用いて、送受信プロセッサユニットを指定し、該送受信プロセッサユニット間でパケットを送受信する方法が採られている。

【０００４】

【発明が解決しようとする課題】上記した従来技術では、パケット受信動作中、または受信待ち動作中に当該プロセッサユニットで障害が発生した場合や、定期保守のために所定のプロセッサユニットをネットワーク網から切り離す必要が生じた場合、処理中のタスクを一旦中断し、プロセッサユニットの切離し、再構成が完了した後に、送信元プロセッサユニットのＯＳ（オペレーティングシステム）は、ネットワーク構成情報を変更し、該変更された物理ネットワークアドレスを用いて、処理中の通信処理を最初からやりなおす必要があり、ＯＳのオーバヘッドが大きくなり、また、通信処理制御が複雑になるという問題があった。

【０００５】本発明の目的は、上記従来技術の問題点を

解決し、必要に応じてプロセッサユニットの動的再構成を高速かつ効率的に行うことのできる並列プロセッサシステムの通信制御方法を提供することにある。

【0006】

【課題を解決するための手段】上記目的を達成するために、本発明は、複数のプロセッサユニットが各々ネットワークルータを介してネットワーク網で結合され、プロセッサユニット間のデータ転送がパケット送受信により行われる並列プロセッサシステムにおいて、システムの状態を監視する監視プロセッサ、あるいは、複数のプロセッサユニットの各ネットワークルータに、各プロセッサユニットの論理アドレスと物理アドレスとの対応を登録したプロセッサアドレス変換テーブルを用意し、パケット送信元プロセッサユニットは宛先プロセッサユニットアドレスとして論理アドレスを付加したパケットを送信し、該パケットを受信したネットワークルータは、前記監視プロセッサに要求を出して前記論理アドレスに対応する物理アドレスを得、該物理アドレスにもとづいてネットワーク網の経路選択を行うようにしたことである。

【0007】

【作用】ネットワーク網から、例えば障害発生プロセッサユニットを切り離す必要が生じた場合、プロセッサアドレス変換テーブル内の当該プロセッサユニットの論理アドレスと対の物理アドレスを代替先プロセッサユニットの物理アドレスに変更する。これにより、障害発生プロセッサユニットのネットワーク網からの切り離しが達成されたことになる。

【0008】ネットワークルータでは、障害発生プロセッサユニットの論理アドレスが付加されたパケットを受信すると、プロセッサアドレス変換テーブルより代替先プロセッサユニットの物理アドレスを得、動的に該代替先プロセッサに対応する経路選択が実施される。また、送信元プロセッサユニットでは、パケットを再送信する場合、代替先プロセッサをまったく意識することなく、送信先プロセッサユニットアドレスとして障害発生プロセッサユニットの論理アドレスが付加されたパケットをそのまま再送信するだけでよい。

【0009】このように、本発明によれば、再構成時のソフトウェア処理が容易になり、ＯＳのオーバヘッドを最小限に抑えることができる。また、高速にネットワーク網からの対象プロセッサの切離し、処理中タスクの継続実行が可能になる。

【0010】

【実施例】以下、本発明の一実施例を図面を用いて具体的に説明する。図１は、本発明の一実施例のブロック構成図である。図において、１０１、１０３、１０５、１０７はプロセッサユニットであり、各々、ネットワークルータ１０２、１０４、１０６、１０８を介し、ネットワーク網を構成するＸおよびＹのクロスバスイッチ１１０、１１１に接続され、並列プロセッサシステムを構成している。１０９は受信パケット中の送信元プロセッサユニットの論理アドレスを一時保持する受信アドレステーブルで、便宜上、図１ではネットワークルータ１０８についてのみ示したが、他のネットワークルータ１０２、１０４、１０６にも同様に具備される。１２０はシステムのネットワーク網の状態を監視するサービスプロセッサ（ネットワーク監視プロセッサ）であり、保守・診断用等の専用の信号線１４０により、プロセッサユニット１０１、１０３、１０５、１０７及びネットワークルータ１０２、１０４、１０６、１０８と接続されている。１２２はサービスプロセッサ１２０が管理するプロセッサアドレス変換テーブルであり、各プロセッサユニット１０１、１０３、１０５、１０７の仮想的なアドレス（論理プロセッサアドレス）とネットワーク網上の物理的なアドレス（物理プロセッサアドレス）との対応を保持している。１３０はシステムコンソールであり、オペレータは該コンソール１３０を使用してプロセッサアドレス変換テーブル１２２の内容を設定することができる。

【0011】なお、ネットワーク網は、通常、Ｘ方向、Ｙ方向、Ｚ方向のクロスバスイッチで構成されているが、図１では説明を簡単にするためにＸ方向、Ｙ方向の二次元のクロスバスイッチのみを示す。また、図１では４台のプロセッサユニットのみが示されているが、前述したように、この種の並列プロセッサシステムは数百から数万のプロセッサユニットで構成されている。

【0012】プロセッサユニット間のデータ転送はパケットの送受信によって行われる。いま、プロセッサユニット１０１を送信元、プロセッサユニット１０７を受信先として、まず、プロセッサユニット１０７が正常の場合の動作を説明する。図４に、この場合のネットワークルータの処理フローチャートを示す。

【0013】ここで、便宜上、プロセッサユニット１０１、ネットワークルータ１０２の座標位置を原点とし、右方向をＸ方向、下方向をＹ方向とする。従って、プロセッサユニット１０１、ネットワークルータ１０２の座標は（０，０）、プロセッサユニット１０３、ネットワークルータ１０４の座標は（０，１）、プロセッサユニット１０５、ネットワークルータ１０６の座標は（１，０）、プロセッサユニット１０７、ネットワークルータ１０８の座標は（１，１）となる。これらの座標値を、各プロセッサユニットの物理プロセッサアドレスとする。

【0014】送信元プロセッサユニット１０１は、ヘッダ部に自プロセッサユニットの論理プロセッサアドレス（送信元論理アドレス）と受信先プロセッサユニット１０７の論理プロセッサアドレス（宛先論理アドレス）が付加されたパケットを送信する。ネットワークルータ１０２は、該パケットを受信すると、ヘッダ部から宛先論

理アドレスを読み出し（ステップ４０１）、信号線１４０を介してサービスプロセッサ１２０へ、該宛先論理アドレスの付加された物理アドレス取得要求を発行することで（ステップ４０２）、サービスプロセッサ１２０から同じく信号線１４０を介して、該宛先論理アドレスに対応する物理アドレス（宛先物理アドレス）を得る（ステップ４０３）。サービスプロセッサ１２０でのプロセッサアドレス変換テーブル１２２による論理／物理アドレス変換動作については、図２により後述する。ネットワークルータ１０２は、自物理アドレスすなわちプロセッサユニット１０１の物理アドレスと取得した宛先論理アドレスとを比較し（ステップ４０４）、同一アドレスか否か判定し（ステップ４０５）、本例では同一でないので、経路選択アルゴリズムで所定のクロスバを選択してパケットを送信する（ステップ４０６）。具体的には、ネットワークルータ１０２は、サービスプロセッサ１２０から宛先物理アドレスとして（１，１）を得て、自分の物理アドレス（０，０）と該宛先物理アドレス（１，１）とを比較し、その差分方向にパケットを送る。ここでは、Ｙ方向のクロスバ１１１を選択してパケットを送るとする（もちろん、Ｘ方向を選択してもよく、これは経路選択アルゴリズムによる）。これにより、次にパケットはネットワークルータ１０４で受信される。

【００１５】パケットを受信したネットワークルータ１０４も、同様にしてサービスプロセッサ１２０から宛先物理アドレス（１，１）を得て、自分の物理アドレス（０，１）と宛先物理アドレス（１，１）とを比較し、一致しないため、その差分方向、つまりここではＸ方向のクロスバ１１２を選択してパケットを送る。

【００１６】このようにして、プロセッサユニット１０１が送信したパケットは、１１２の経路でネットワークルータ１０８まで到達する。ネットワークルータ１０８でも、パケットを受信すると、同様にサービスプロセッサ１２０から宛先物理アドレス（１，１）を得て、自分の物理アドレス（１，１）と比較する。この結果、アドレスが一致するので、該ネットワークルータ１０８は、受信パケットを自プロセッサユニット１０７に送る（ステップ４０７）。この時、ネットワークルータ１０８は、受信パケットに付加されている送信元論理アドレスを受信アドレステーブル１０９に一時登録しておき、受信パケットがすべて正常にプロセッサユニット１０７に取り込まれると、当該送信元論理アドレスを削除する。この受信アドレステーブル１０９の詳細は図３により後述する。

【００１７】次に、受信先プロセッサユニット１０７で障害が発生した場合について説明する。プロセッサユニット１０７は、障害が発生すると、信号線１４０を介してサービスプロセッサ１２０へ障害発生割込みを通知する。サービスプロセッサ１２０は、システムをフリーズ

（凍結）した後、プロセッサアドレス変換テーブル１２２を書き替え、プロセッサユニット１０７の論理プロセッサアドレスに対応する物理プロセッサアドレスを変更し、フリーズを解除する。ここでは、プロセッサユニット１０５の物理プロセッサアドレス（１，０）に変更されたとする。

【００１８】いま、フリーズ時、プロセッサユニット１０１の送信したパケットがネットワークルータ１０４とネットワークルータ１０８の途中のクロスバ上にあったとすると、フリーズ解除で、該パケットはネットワークルータ１０８に到達する。ネットワークルータ１０８は、パケットを受信すると、上記と同様にしてサービスプロセッサ１２０から宛先物理アドレスを取得するが、この時、サービスプロセッサ１２０からは宛先物理アドレスとしてプロセッサユニット（代替受信先プロセッサユニット）１０５の物理アドレス（１，０）が返ってくる。ネットワークルータ１０８は、自分の物理アドレス（１，１）と宛先物理アドレス（１，０）とを比較し、一致しないため、その差分方向、即ち、ここではＹ方向のクロスバ１１１を選択してパケットを送信する。この結果、１１３の経路で、パケットは代替受信先プロセッサユニット１０５のネットワークルータ１０６に受信される。ネットワークルータ１０６は、サービスプロセッサ１２０から同様に宛先物理アドレス（１，０）を得、自分の物理アドレス（１，０）と一致するため、受信パケットを自プロセッサユニット１０５へ取り込む。

【００１９】図２は、サービスプロセッサ１２０が管理するプロセッサアドレス変換テーブル１２２の具体的構成例、及び、該変換テーブル１２２を用いたプロセッサユニット間の通信および動的な再構成を説明する図である。

【００２０】図１で説明したと同様にして、送信元プロセッサユニット２０１から送信されたパケットは、ネットワークルータ、クロスバスイッチを経由して受信先プロセッサユニット２０３に転送される。そのとき、各ネットワークルータは、パケットのヘッダ部に付加された宛先論理アドレスに対応する物理アドレスの取得要求をサービスプロセッサ１２０へ発行する。これを受けて、サービスプロセッサ１２０では、テーブル参照回路１２１を介してプロセッサアドレス変換テーブル１２２を参照し、宛先論理アドレス（物理プロセッサアドレス）に対応する物理アドレス（物理プロセッサアドレス）を読み出して、要求のあったネットワークルータへ返送する。

【００２１】プロセッサアドレス変換テーブル１２２の各エントリは、プロセッサグループのクラスを指定するクラスビット、エントリの有効無効を判断する有効ビット、論理プロセッサアドレス、及び、論理プロセッサアドレスに対応する物理プロセッサアドレスで構成される。該プロセッサアドレス変換テーブル１２２は、サー

ビスプロセッサ１２０が発行するＤＩＡＧ（診断）命令により初期化され、障害が発生した場合は、サービスプロセッサ１２０により自動的に、又は、システムコンソール１３０からマニアル指示により書替えられる。

【００２２】クラスビットは、サービスプロセッサ１２０のＯＳが初期化する際に、グループ分けしたプロセッサユニット群毎にシリアル番号を持っている。例えば、ｎ台のプロセッサユニットからなる並列プロセッサシステムを、２つのグループに論理的に分けて動作させたとき、一方のプロセッサグループに例えばクラス００１を割り当て、他方のプロセッサグループにクラス００２を割り当てる。有効ビットは、処理参加中のプロセッサユニットに対応するエントリには“１１”（使用中）が設定され、論理的に切り離されているが、動作可能なプロセッサユニットに対応するエントリには“１０”（待期中）が設定され、論理的に切り離されていて、且つ動作不可であるプロセッサユニットに対応するエントリには“００”（無効）が設定されている。論理プロセッサアドレスはプロセッサユニットに任意に割り当てた仮想アドレス、物理プロセッサアドレスはネットワーク網の実際の物理アドレスを示す。

【００２３】いま、送信元プロセッサユニット２０１の論理プロセッサアドレスを“００００００００１”、その物理プロセッサアドレスを“００００００００”とし、受信先プロセッサユニット２０３の論理プロセッサアドレスを“００００００００９”、その物理プロセッサアドレスを“０００２０００２”とする。また、代替受信先プロセッサユニット２０５の論理プロセッサアドレスを“００００００００４”、その物理プロセッサアドレスを“０００１０００”とする。

【００２４】送信元プロセッサユニット２０１から受信先プロセッサユニット２０３へデータ転送する場合、プロセッサユニット２０１は、送信元論理アドレスとして論理プロセッサアドレス“００００００００１”、宛先論理アドレスとして論理プロセッサアドレス“０００００００９”を付加したパケットを送信する。該パケットを受信した各ネットワークルータは、宛先論理アドレス“００００００００９”を指定して、それに対応する物理アドレスの取得要求を、サービスプロセッサ１２０へ発行する。サービスプロセッサ１２０のテーブル参照回路１２１は、プロセッサアドレス変換テーブル１２２より論理プロセッサアドレス“００００００００９”に対応する論理プロセッサアドレス“０００２０００２”を検索し、要求のあったネットワークルータへ返送する。このようにして、受信先プロセッサユニット２０３が正常であれば、プロセッサユニット２０１から送信されたパケットは、経路２１０でプロセッサユニット２０３に受信される。

【００２５】一方、障害発生時におけるパケットルーチングの動的再構成は、プロセッサアドレス変換テーブル１２２により以下のようにして行う。まず、プロセッサアドレス変換テーブル１２２について、障害プロセッサユニットと同一クラス内で有効ビットが“１０”となっているエントリに対応するプロセッサユニットを、代替受信プロセッサユニットとして選択する。次に、障害発生プロセッサユニットの論理プロセッサアドレスに対応する物理プロセッサアドレスを、該選択されたプロセッサユニットの物理プロセッサアドレスに書き換え、該選択されたエントリの有効ビットを“１０”から“００”に書き換えて無効にする。

【００２６】例えば、受信先プロセッサユニット２０３に障害が発生すると、該プロセッサユニット２０３は、自論理プロセッサアドレス“００００００００９”を指定して、障害発生割込みをサービスプロセッサ１２０へ通知する。サービスプロセッサ１２０は、プロセッサアドレス変換テーブル１２２を参照して、同一クラス（００１）内で有効ビットが“１０”となっているエントリとして、論理プロセッサアドレスが“００００００００４”で、物理プロセッサアドレスが“０００１０００”であるプロセッサユニット２０５を代替受信プロセッサユニットとして選択する。そして、障害が発生したプロセッサユニット２０３のエントリの論理プロセッサアドレス“００００００００９”に対応する物理プロセッサアドレス“０００２０００２”を“０００１０００”に書き換えるとともに、選択されたエントリの有効ビット“１０”を“００”に書き換えて無効化する。この結果、フリーズ解除後、パケットを受信したネットワークルータ２０４が、宛先論理アドレス“０００００００９”を指定して物理アドレス取得要求をサービスプロセッサ１２０へ発行すると、サービスプロセッサ１２０のテーブル参照回路１２１は、プロセッサアドレス変換テーブル１２２より論理アドレス“００００００００９”に対応する論理プロセッサアドレス“０００１０００”を検索してネットワークルータ２０４へ返送する。これにより、経路２１１を経由して、パケットが代替受信先プロセッサユニット２０５で受信可能となる。

【００２７】図３は、各ネットワークルータが具備する受信アドレステーブルの具体的構成例、及び、その働きを説明する図である。

【００２８】図３において、３０１、３０３を送信元プロセッサユニット、３０５を受信先プロセッサユニットとする。３０６は受信先プロセッサユニット３０５のネットワークルータであり、テーブル参照回路３１０、受信アドレステーブル３１１を具備する。受信アドレステーブル３１１のエントリは、当該エントリの有効・無効を示すＶＬＤビットと、受信パケットの送信元論理アドレスが登録される論理アドレスとで構成される。図３では、ネットワークルータ３０６についてのみ示したが、このような受信アドレステーブルが、すべてのネットワークルータ１０２に内蔵されている。

【００２９】いま、送信元プロセッサユニット３０１から送信されたパケットが受信先プロセッサユニット３０５のネットワークルータ３０６に到達すると、当該ネットワークルータ３０６では、テーブル参照回路３１０を介して、受信アドレステーブル３１１に、受信パケットに付加されている送信元プロセッサユニット３０１の論理アドレス（送信元論理アドレス）を登録し、そのエントリのＶＬＤビットに“１”を設定する。そして、当該送信元プロセッサユニット３０１からのパケットがすべてプロセッサユニット３０５に取り込まれると（受信完了）、該ネットワークルータ３０６は、該当エントリのＶＬＤビットを“０”に書き換える。テーブル参照回路３１０は、１エントリの登録が終了すると、エントリのポインタを次のエントリアドレスにポイントし、受信アドレステーブル３０６がフルになった場合、ＶＬＤビットが“０”となっている低位アドレスのエントリの先頭をポイントする。

【００３０】ここで、受信先プロセッサユニット３０５で障害が発生した場合、該プロセッサユニット３０５はサービスプロセッサに障害発生割込みを通知する。これを受けてサービスプロセッサでは、図２で説明したようにプロセッサアドレス変換テーブルを書き替えてパケットルーティングの動的再構成を行う。その後、サービスプロセッサは、該障害発生プロセッサユニット３０５に接続されたネットワークルータ３０６内の受信アドレステーブル３１１の内容をテーブル参照回路３１０を通して読み取り、ＶＬＤビットが“１”すなわちパケットの受信が完了していない論理アドレスに対応する送信元プロセッサユニット（図３では３０１）に対してパケットの再送要求割込みを報告する。

【００３１】パケットの再送要求に対して、送信元プロセッサユニット３０１のＯＳでは、受信先プロセッサユニット３０５の障害発生をまったく意識することなく、障害発生前と同じ条件設定のままで、パケットの送出指示を行う。この結果、該送信元プロセッサユニット３０１から再送されたパケットは、図２で説明したようにして、所定の代替受信先プロセッサユニットで受信される。

【００３２】一方、図３において、他の送信元プロセッサユニット３０３からのパケットは、まだクロスバスイッチ上にあり、受信先プロセッサユニット３０５に到達していないので、ネットワークルータ３０６の受信アドレステーブル３１１にはその論理アドレスは登録されていない。したがって、サービスプロセッサからは該送信元プロセッサユニット３０３に対して、パケットの再送要求割込みを報告されない。これは、途中のパケットは、フリーズ解除後、自動的に代替受信先プロセッサユニットで受信されることによる。

【００３３】また、受信アドレステーブル３１１内において、ＶＬＤビットが“０”、すなわち、障害発生プロセッサユニット３０５で障害発生前にパケットの受信が完了した論理アドレスに対応する送信元プロセッサユニットに対しても、サービスプロセッサからはパケットの再送要求割込みを報告されない。これは、図４で後述するように、サービスプロセッサが障害発生プロセッサユニットから代替受信先プロセッサユニットにコピーする情報に正常なパケット内容が含まれていることによる。

【００３４】図５は、任意プロセッサユニットで障害が発生してから代替プロセッサユニットで処理の継続を開始するまでの、サービスプロセッサの全体的な処理フローチャートを示したものである。

【００３５】ある受信先プロセッサユニットで障害が発生すると、該プロセッサユニットは障害発生割込みをサービスプロセッサに対して報告する（ステップ５０１）。サービスプロセッサは、障害報告割込みを受け取ると、直ちにシステム全体の動作をフリーズ（凍結）する（ステップ５０２）。次いで、図２で説明したように、プロセッサアドレス変換テーブルを参照して、同一クラス内で代替プロセッサユニットの対象となるエントリ（有効ビットが“１０”のエントリ）を選択し（ステップ５０３）、該変換テーブルに登録されている障害発生受信先プロセッサユニットのエントリの物理プロセッサアドレスを、選択された代替プロセッサユニットのエントリの物理プロセッサアドレスに書き換え（ステップ５０４）、代替プロセッサユニットのエントリを無効化（有効ビットを“００”）する（ステップ５０５）。その後、障害が発生した受信先プロセッサユニットのハードウェア情報を、代替受信先プロセッサユニットにコピーする（ステップ５０６）。

【００３６】次いで、サービスプロセッサは障害が発した受信先プロセッサユニットに接続されているネットワークルータ内の受信アドレステーブルを読み込み（ステップ５０７）、ＶＬＤビットにより、該障害発生受信先プロセッサユニットがパケット受信処理中かどうか判定する（ステップ５０８）。そして、障害発生受信先プロセッサユニットがパケット受信処理中（ＶＬＤビットが“１”）の場合、受信アドレステーブルの当該エントリに登録されている論理アドレスの送信元プロセッサユニットに対して、パケットの再送要求割込みを報告し（ステップ５０９）、代替受信先プロセッサユニットに対しては、ステップ５０６でコピーした障害発生プロセッサユニットのハードウェア情報中の受信中パケットの破棄割込みを報告して（ステップ５１０）、システムフリーズを解除する（ステップ５１１）。また、障害発生プロセッサユニットがパケット受信処理中でない場合は（ＶＬＤビットが“０”）、ステップ５０９、５１０の処理を行うことなく、システムフリーズを解除する。

【００３７】図６は、本発明の第２の実施例のブロック構成図で、各ネットワークルータ６０２、６０４、６０６、６０８にそれぞれプロセッサアドレス変換テーブル

６２２１～６２２４を用意したものである。なお、図６では省略したが、該変換テーブルの他に、各ネットワークルータ６０２、６０４、６０６、６０８は図３で説明した受信アドレステーブルを持つことは第１の実施例と同様である。

【００３８】サービスプロセッサ６２０はマスタプロセッサアドレス変換テーブル６２２を持っており、システム立上げ時、該マスタプロセッサアドレス変換テーブル６２２の内容を信号線６４０により各ネットワークルータ６０２、６０４、６０６、６０８に配布してプロセッサアドレス変換テーブル６２２１～６２２４を初期設定する。また、任意プロセッサユニットで障害が発生した場合、サービスプロセッサ６２０は、図２で説明したと同様にしてマスタプロセッサアドレス変換テーブル６２２を書き換え、その変更後のマスタプロセッサアドレス変換テーブル６２２の内容を、信号線６４０を介して、各ネットワークルータ６０２、６０４、６０６、６０８のプロセッサアドレス変換テーブル６２２１～６２２４に再設定する。これにより、各ネットワークルータ内のプロセッサアドレス変換テーブルの一致性が保証される。

【００３９】図６の構成によれば、各ネットワークルータ６０２、６０４、６０６、６０８はパケットを受信した時、それぞれ自分のアドレス変換テーブル６２２１、６２２２、６２２３、６２２４を参照して宛先論理アドレスに対応する物理アドレスを得ることができ、一々、サービスプロセッサ６２０へ物理アドレス取得要求を発行する必要がなくなる。また、サービスプロセッサ６２０においても、該物理アドレス取得要求に一々応答する必要がなくなる。

【００４０】以上、本発明の一実施例について説明したが、パケット転送路としては、クロスバスイッチに限定されるものではなく、他の任意のものを用いることができる。また、並列プロセッサシステムを構成するプロセッサユニット群の１台あるいは複数台のプロセッサユニットにサービスプロセッサの機能を兼ねさせ、サービスプロセッサが故障した場合、当該プロセッサユニットにサービスプロセッサの処理を代替させるようにしてもよい。

【００４１】

【発明の効果】

（１）本発明によれば、ネットワーク網に接続されている全てのプロセッサユニットを論理アドレスと物理アドレスで管理しているので、ネットワーク網から対象プロセッサユニットを容易に切離すことができ、動的な再構成処理を高速に行うことができ、可用性、保守性が格段に向上する。

【００４２】（２）各プロセッサユニットの論理アドレスと物理アドレスの対応を登録したプロセッサアドレス変換テーブルは、システムの状態を監視する監視プロセッサでの一元管理、あるいは、各プロセッサユニット対応のネットワークルータでの分散管理のいずれでもよい。ここで、監視プロセッサでの一元管理は、保守が簡単であるが、ネットワークルータが一々、監視プロセッサに参加要求を発行する必要があるため、プロセッサユニットの構成数があまり多くない場合に向いている。一方、各ネットワークルータでの分散管理は、監視プロセッサへの参照要求が不要であり、プロセッサの構成数が多量の場合に有効である。

【００４３】（３）各ネットワークルータは、対応するプロセッサユニットがパケット受信処理中であるか否かを示す識別ビット、及び、当該パケットの送信元プロセッサユニットの論理アドレスの対応を管理しているので、受信中に障害があった送信元プロセッサユニットに対してのみ再送信処理を行い、パケット送信途中で、まだ受信側に到達していないパケットは、動的に再構成されても再送信処理を行うことなく、新たに割り付けられた受信先プロセッサユニットに送信される。従って、プロセッサユニットのＯＳの再送信処理が簡素化され、処理オーバヘッドを大幅に削減することができる。

【図面の簡単な説明】

【図１】本発明の第１の実施例のブロック構成図である。

【図２】プロセッサアドレス変換テーブルの構成例およびネットワークルーチンの動的な再構成を説明する図である。

【図３】受信アドレステーブルの構成例およびその働きを説明する図である。

【図４】ネットワークルータのパケット受信時の処理フローチャートである。

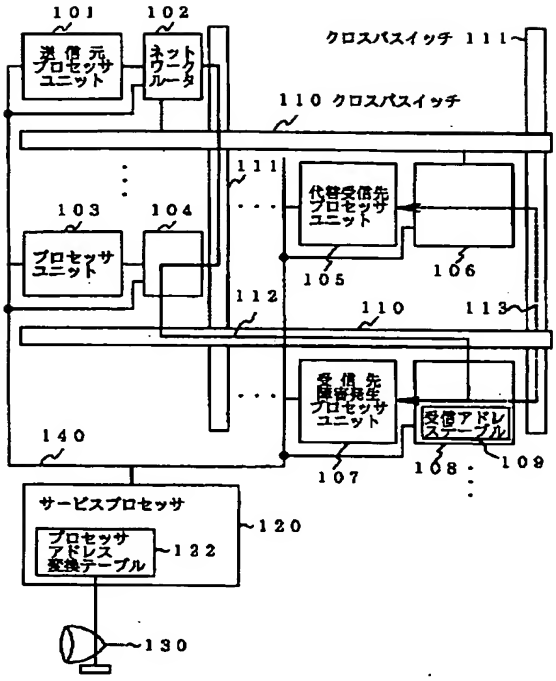
【図５】障害発生時のサービスプロセッサの動的再構成制御の処理フローチャートである。

【図６】本発明の第２の実施例のブロック構成図である。

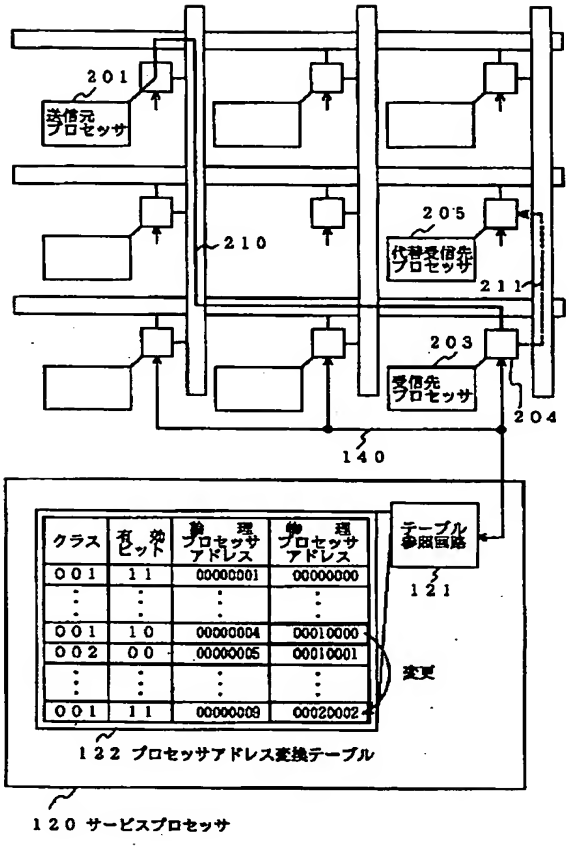
【符号の説明】

１０１、１０３、１０５、１０７　プロセッサユニット
１０２、１０４、１０６、１０８　ネットワークルータ
１０９　受信アドレステーブル
１１０、１１１　クロスバスイッチ
１２０　サービスプロセッサ
１２２　プロセッサアドレス変換テーブル

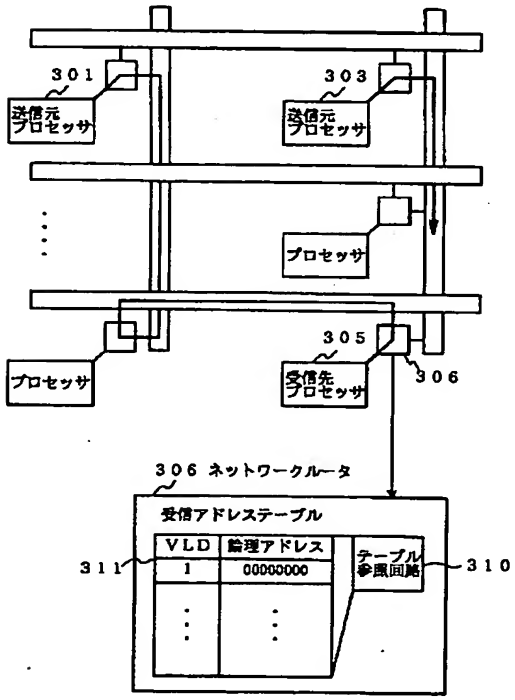
【図1】


101
102
送信元プロセッサユニット
ネットワークルータ
クロスバスイッチ 111
110 クロスバスイッチ
111
代替受信先プロセッサユニット
103
104
プロセッサユニット
105
106
112
110
113
受信先障害発生プロセッサユニット
受信アドレステーブル
140
107
108
109
サービスプロセッサ
プロセッサアドレス変換テーブル
120
122
130

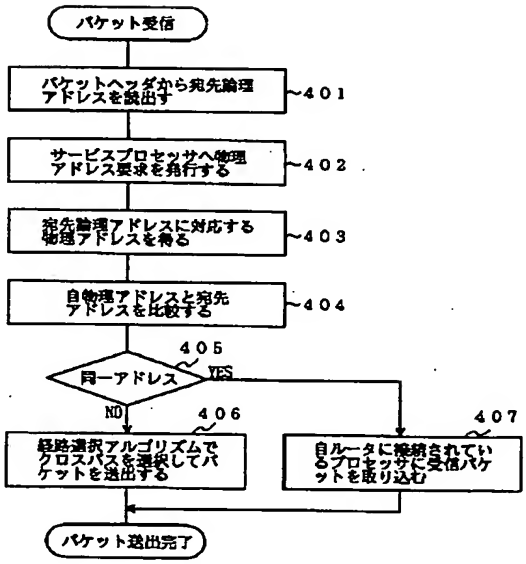
【図2】


201
送信元プロセッサ
205
代替受信先プロセッサ
210
211
203
受信先プロセッサ
204
140
クラス
有効ビット
論理プロセッサアドレス
物理プロセッサアドレス
テーブル参照回路
121
001 11 00000001 00000000
001 10 00000004 00010000
002 00 00000005 00010001
001 11 00000009 00020002
変更
122 プロセッサアドレス変換テーブル
120 サービスプロセッサ

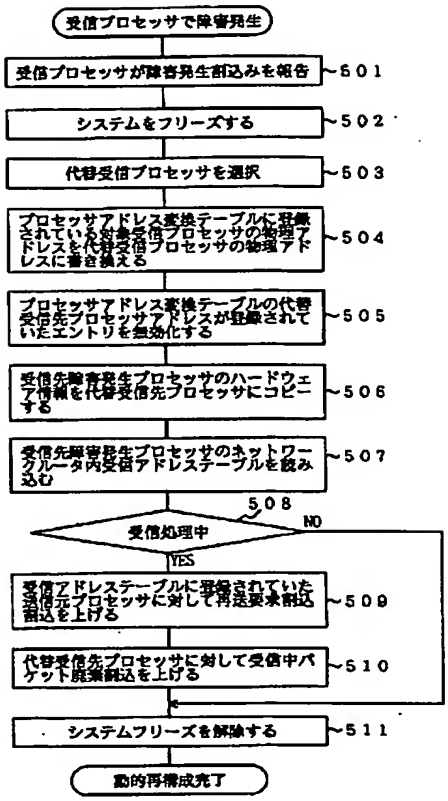
【図3】


301
303
送信元プロセッサ
送信元プロセッサ
プロセッサ
プロセッサ
305
306
受信先プロセッサ
306 ネットワークルータ
受信アドレステーブル
VLD
論理アドレス
1
00000000
311
テーブル参照回路
310

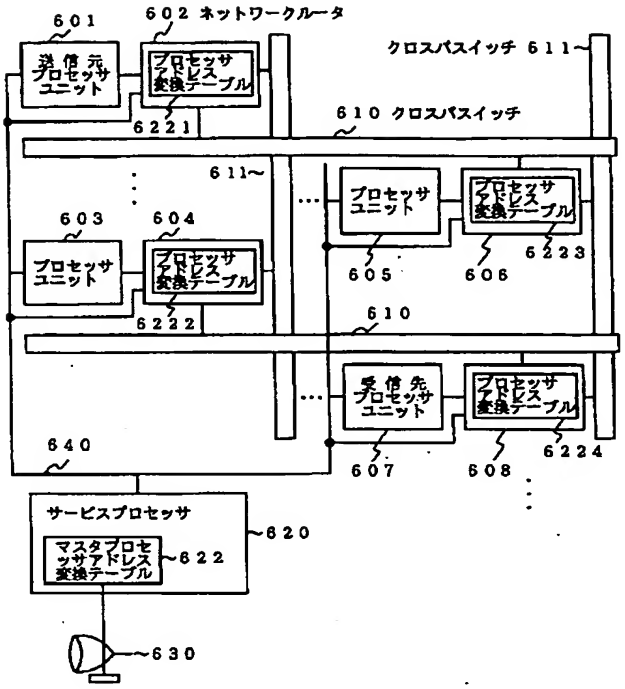
【図4】


パケット受信
パケットヘッダから宛先論理アドレスを抽出す
401
サービスプロセッサへ物理アドレス要求を発行する
402
宛先論理アドレスに対応する物理アドレスを得る
403
自物理アドレスと宛先アドレスを比較する
404
405
同一アドレス
YES
NO
406
経路選択アルゴリズムでクロスバスを選択してパケットを送出する
407
自ルータに接続されているプロセッサに受信パケットを取り込む
パケット送出完了

【図５】


受信プロセッサで障害発生
受信プロセッサが障害発生割込みを報告
501
システムをフリーズする
502
代替受信プロセッサを選択
503
プロセッサアドレス変換テーブルに登録されている対象受信プロセッサの物理アドレスを代替受信プロセッサの物理アドレスに書き換える
504
プロセッサアドレス変換テーブルの代替受信先プロセッサアドレスが登録されていたエントリを無効化する
505
受信先障害発生プロセッサのハードウェア情報を代替受信先プロセッサにコピーする
506
受信先障害発生プロセッサのネットワークルータ内受信アドレステーブルを読み込む
507
508
受信処理中
NO
YES
受信アドレステーブルに登録されていた送信元プロセッサに対して再送要求割込割込を上げる
509
代替受信先プロセッサに対して受信中パケット廃棄割込を上げる
510
システムフリーズを解除する
511
動的再構成完了

【図６】


601
602 ネットワークルータ
送信元プロセッサユニット
プロセッサアドレス変換テーブル
6221
クロスバスイッチ 611
610 クロスバスイッチ
611
603
604
プロセッサユニット
プロセッサアドレス変換テーブル
6222
605
606
6223
610
受信先プロセッサユニット
607
608
6224
640
サービスプロセッサ
マスタプロセッサアドレス変換テーブル
622
620
630